SMB

β-Hexosaminidase and a DNA Sequence Coding it Obtained from Ciliates and Use Thereof

The present invention relates to a nucleic acid coding for a β-hexosaminidase.

The expression of foreign proteins in microorganisms, such as bacteria, yeasts or mammal cells, is of great importance to the biotechnological preparation and production of recombinant proteins. Thus, bacterial expression systems based on *E. coli* or *B. subtilis* are used for the production of recombinant peptides or proteins, such as insulin, interleukin-2, tissue plasminogen activator, proteases and lipases. In Gram-negative bacteria, the expression systems are mostly based on the use of genetic elements such as the lac operon or the tryptophan operon. The proteins foreign to the host are produced either into "inclusion bodies" within the cell, or when expression systems based on β-lactamase genes are used, into the periplasmic space. The production of recombinant proteins into the surrounding fermentation medium has not been established. In Gram-Positive bacteria, to date, almost exclusively cell-inherent proteins are introduced and expressed.

Yeasts, such as *S. cerevisiae*, *Kluyveromyces lactis* or *Pichia pastoris*, are also employed for the heterologous expression of recombinant proteins, such as human factor XIIIa, bovine pro-chymosin, or surface antigens. In yeasts, the expression systems are based on shuttle vectors (vectors having both yeast and bacterial portions) which are constructed from the genetic elements of galacto-kinase-epimerase, acid phosphatase or alcohol-dehydrogenase. As a rule, the recombinant protein is produced into the cytoplasm of the cell. When yeast-inherent signal sequences, such as the alpha factor, are used, the expressed proteins may also be secreted into the fermentation medium. The glycosylation of secreted proteins is effected according to the "high mannose" type.

Mammal cells, such as various cell types from rodents (CHO cells, C127 cells) or simians (vero, CV-1 or COS cells), are also employed for the heterologous expression of recombinant proteins. Here, the expression systems are based on recombinant viruses (BPV vector) or on shuttle vectors. To regulate the expression, viral SV40 enhancer/promoter systems or cellular enhancer elements are employed. The recombinant proteins, such as erythropoietin, are secreted into the fermentation medium because the foreign genes usually bring their own signal sequences, which are understood by the expression system and used for targeting.

Further, for the biotechnological production of glycosylated extracellular enzymes, ciliates such as *Tetrahymena* are employed. Ciliates will grow on inexpensive fermentation media using standard fermentation methods. For the transformation of such ciliates, vectors are available which are based on the rDNA elements of the ciliate *Tetrahymena*. For the heterologous expression of bacterial proteins in ciliates, DNA constructs consisting of genes from *Tetrahymena* are employed. When suitable genetic elements for the regulation of the transcription, targeting and glycosylation of foreign proteins are available, ciliates are an ideal expression system for the inexpensive production of therapeutic recombinant proteins.

The Gram-negative bacterial expression systems used to date usually lead to the formation of "inclusion bodies" in the cell, accompanied by a denaturing of the proteins. To recover the recombinant protein, the cells must be lysed, and the denatured inactive protein must be folded back to function. This causes additional cost-intensive process steps and reduces the yield of the desired protein. Glycosylation, which is important to eukaryotic proteins, is completely omitted. When Gram-positive bacterial expression systems are used, degradation of the target protein due to high proteolytic activities in the fermentation broth is an additional problem.

When yeasts are used for heterologous expression, the desired target protein is often produced into the cell only from where it must be removed by cell lysis. As in bacterial expression systems, this causes additional time- and cost-intensive process steps. When yeast-inherent signal peptides are used, the foreign proteins are not correctly spliced and glycosylated for secretion.

In contrast, when mammal cell systems are employed for the production of recombinant proteins, the desired proteins are found in the fermentation medium in an extracellular state, correctly spliced and glycosylated. However, what is disadvantageous here is, on the one hand, the low expression rate due to the defective processing and inefficient translation of genes which have been introduced into the genome of the production cell line via viral vectors. On the other hand, the serum-containing fermentation media for mammal cells are extremely cost-intensive. In addition, the fermentation technology for the shear-sensitive cell lines is complicated and similarly expensive due to construction for bubble-free aeration. Further problems arise from the high infection risk for the cell lines from mycoplasmas and viruses. All in all, the use of mammal cells for the biotechnological preparation of recombinant proteins results in very high costs, safety demands and low yields.

To the use of ciliates, such as *Tetrahymena*, the above mentioned drawbacks in the production of proteins do not apply. Thus, for example, some acid hydrolases which are involved in the digestion of food particles are exported from the cell in high quantities and with complex glycosylation.

In J. Euk. Microbiol. 43 (4), 1996, pages 295 to 303, Alam et al. describe the cloning of a gene which codes for the acid α-glucosidase of *Tetrahymena pyriformis*. However, only a small portion of the protein is exported from the cell.

However, to date, it has not been possible to bring other foreign glycosylated eukaryotic proteins to expression in ciliates which are also secreted into the fermentation medium. This is due to the fact that the DNA sequences from ciliate-inherent secreted acid hydrolases which are necessary for the construction of expression vectors have not been known to date.

It has been the object of the invention to provide a DNA sequence for the expression of secreted proteins of ciliates. The DNA sequence is to enable heterologous proteins in an expression system to be exported into the fermentation medium after transformation in ciliates. This system is also to achieve a high expression

rate of the heterologous protein, which is exported in large amounts from the cell under culture conditions.

This object is achieved by a system in which a nucleic acid having a sequence with Seq. Id. No. 1 and coding for β-hexosaminidase is used.

In particular, the inventive DNA sequence of the β-hexosaminidase includes a signal peptide and a propeptide, and optionally further genetic elements for the targeting of proteins. The use of these sequences in a vector enables heterologously expressed proteins to be exported from the cell and thus to be purified from the fermentation broth without a cell lysis.

Figure 1 shows a nucleic acid coding for β-hexosaminidase from ciliates. Figure 2 shows a corresponding expression product of the nucleic acid according to Seq. Id. No. 1. The invention also relates to this protein according to Seq. Id. No. 2.

In particular, the invention relates to the signal sequence of the protein according to the invention. This is preferably amino acids 1 to 17 of the protein according to the invention. The invention also relates to a nucleic acid which codes for the N-terminal fragment. This is preferably a fragment of the nucleic acids according to the invention, especially with the nucleic acid sequence from 1 to 51 according to Figure 1.

A further aspect of the invention is the use of a nucleic acid sequence of acid hydrolases according to the invention or parts thereof for the homologous or heterologous expression of recombinant proteins and peptides, and for homologous or heterologous recombination ("knock-out, "gene replacement").

The invention also relates to a method in which the nucleic acid or parts thereof according to the invention which code for β-hexosaminidase is combined with the usual, in homologous or heterologous expression, enhancers, promoters, operators, origins, terminators, antibiotic resistances, or other nucleic acids or DNA fragments, or sequences of any kind from viroids, viruses, bacteria, archezoans, protozoans, fungi, plants, animals or humans.

In particular, the nucleic acid or parts thereof according to the invention is inserted into a vector, a plasmid, a cosmid, a chromosome or minichromosome, a transposon, an IS element, an rDNA, or any kind of circular or linear DNA or RNA.

The skilled person will understand that nucleic acids having at least 40% homology with the nucleic acid according to Seq. Id. No. 1 can also be employed according to the invention. The protein according to Seq. Id. No. 2 can also be modified without losing its function. Thus, for example, so-called conservative exchanges of amino acids may be performed. Thus, for example, hydrophobic amino acids can be interchanged.

For the purification and isolation of β-hexosaminidase from ciliates and for determining its sequence, the following methods can be used. This is illustrated by the following Examples.

Example 1

From a total of 3.2 l of cell culture, cells of the ciliate *Tetrahymena* in late logarithmic growth phase were washed into 400 ml of starving medium (10 mM Tris-HCl, pH 7.4), and the cells were incubated for another 4 hours with shaking. Then, the cell-free culture supernatant was harvested and filtered through a number of filters with decreasing pore diameters to remove any particular material. The filtrate was concentrated with an Amicon ultrafiltration cell and rebuffered into the starting buffer for the subsequent ion-exchange chromatography (IEC).

The collected chromatographic fractions were tested with specific 4-nitrophenyl substrates for acid hydrolases, acid phosphatase (aPAse) and β-hexosaminidase (β-Hex), and the fractions having the highest β-Hex activity were combined. They were concentrated by another Amicon ultrafiltration and rebuffered into the starting buffer for affinity chromatography. The collected chromatographic fractions were tested for enzymatic activities as described above, and the fractions having the highest activity for an acid hydrolase were combined and rebuffered into phosphate buffer (PB). From a total of eight purifications, the chromatographic

fractions containing hydrolases were pooled, concentrated, portioned and frozen until further characterization.

Part of the thus purified hydrolase was separated by two-dimensional SDS gel electrophoresis (a total of eight gels), the main spot was respectively punched out and collected. The protein present in these gel pieces, β-Hex, was digested with the protease trypsin. The thus obtained peptide fragments were leached out of the gel pieces and separated by reverse-phase HPLC (RP-HPLC). The purity of the HPLC fractions was tested with a mass-spectroscopical method (MALDI-MS), and the pure fractions, i.e., those which contained only one peptide species of a defined mass, were sequenced from their N-terminus using Edman degradation.

From the sequence of the peptide fragments obtained by the trypsin digestion, β-hexosaminidase-specific PCR primers were established. By a sequence comparison with the existing β-Hex sequences, a preselection of the primer combinations for RT-PCR could be made. With this information, the first specific cDNA fragments were amplified and sequenced, using RT-PCR, from isolated whole RNA whose quality was previously tested by Northern hybridization. Using these fragments, sequence-specific primers were constructed which were employed for further PCR experiments. Each primer and each partial sequence obtained was checked by data base alignment, and prior to further experiments, it was also checked for any overlooked vector sequences, inter alia. By 5'- and 3'-RACE, the cDNA sequence was elongated at the 5'- and 3'-termini. The thus completed sequence of the β-Hex cDNA then served as the basis for further sequence analyses.

The thus established sequence of a β-hexosaminidase from the ciliate reads as set forth in Figure 1. In total, this sequence comprises 1836 base pairs including 5'- and 3'-non-translated regions of β-hexosaminidase with an open reading frame having a length of 1656 base pairs. The complementary amino acid sequence has a length of 551 amino acids and reads as set forth in Figure 2.

The sequence of the β-hexosaminidase has a total of 9 glycosylation sites and contains a signal peptide and a pro sequence for targeting the enzyme through the sorting mechanism of the cell.